要因実験における因子の種類と推測の方法(実験データ解析の理論的背景)

竹内, 啓

数理解析研究所講究録
526巻 1984年 1-12

要因実験における因子の種類と推測の方法

竹内 啓

実験計画法における基礎概念の一つに、因子効果の種別がある。これは主として日本の統計家によって発展されたものであるといえる。外国の文献にも勿論それは見られるが、奥野忠一氏や田口玄一氏によってなされた議論によって、精密化されたもので、まだ外国にはあまり知られていないが、わが国の應用統計学の重要な成果であるということができよう。しかしその種別は主として計画の段階で用いられ、データ解析、推測の段階には十分反映されていないように思われる。けれども、それは適切な推測方式を定めるに際しても重要な意味を持つのである。以下その主要な点についてのべよう。

因子についてふつうに用いられる区別は、次のようなものである。

a) 制御因子。その水準を制御することによって、望ましい結果を得るために用いられる因子。

b) 標示因子。実験条件に影響を与える因子であって、かつその水準が再現性を持つもの。

c) ブロック因子。実験条件に影響を与える因子で、その水準は定義可能であるが、再現性を持たないもの

d) 補助因子。実験条件に影響を与える因子で、その水準は

再現性を持つが，事前には認知できないもの

e) 潜在因子．実験条件に傾向を持った影響を与えるが，認知不能なもの，あるいは見落されているもの．これは本来はあってはならないものである．

f) 誤差．その変動が偶然的と思われるすべての因子の影響をまとめたもの．

ここでは制御因子，標示因子，ブロック因子の3種類について主としてのべよう．ただし最初に断っておくべきことは，このような区別は絶対的なものではなく，問題意識と，状況に応じて変わるものであるという点である．例えばある種の農作物の収量を比較する実験において，最適な品種と肥料の組み合わせを求めることを考えるとすれば，品種も肥料もともに制御因子となるが，いろいろな理由からえらばれたそれぞれの品種に対して，最適な肥料を求めることにすれば，品種は標示因子となる．またこのような実験が何か所かで行われる場合，ふつう場所は標示因子と考えられるが，もし最適地を探すということが実験の目的であれば，それは制御因子になる．もう一つ注意すべきことは，上記の因子の区別は実験の場においてではなく，現実の適用の場を前提にしてなされねばならないことである．すなわち標示因子とは，一般に応用の場において再現性があり，かつ制御不可能なもの，ブロ

つ7因子とは応用の場において再現性を持たないものを意味するのである。

すべての実験においては，少くとも一つの制御因子がふくまれなければならない。そうでなければ実験を行う意味がない。しかしその他の因子の組み合わせについてはいろいろな可能性がある。その組み合わせの如何によって，モデルにも解析法にも差が出るはずである。

まずモデルについて考えよう。ここで因子水準はすべて離散的，反応は連続量と仮定しておく。

この場合因子効果について母数模型（fixed-effect model）と変量模型（random effect model）の2種類があることはよく知られている。しかし本来制御因子や標示因子についてその効果を確率的な変量と仮定することは妥当でない。制御因子の効果を変量とすることは，概念上矛盾であるし また標示因子においては，その効果が再現性のない確率変量であるならば それは本来ブロック因子と考えねばならない。しかしこのように考えると，ブロック因子の概念を，ふつうに理解されているより，やや拡大して解釈しなければならないことになる。というのは応用の場においては再現性がなく，制御もできないが，実験の場においては再現性を持っている，

或いは制御可能であるような因子もあるが，このような場合においては実験の場でえらばれた因子水準が，応用の場における因子水準の集団からランダムにえらばれたものであると見なし得る限り，その効果を確率変量と見ることができる。そうするとこのような因子は本来ブロック因子と見なすべきではあるが，実験の場のみに注目する限り標示因子と見ることも可能であり，また実際にブロックとは見なされないことが多い。そうしてその因子効果のばらつきの大きさを知ることが重要な意味を持つ場合も少くない。薬の効果の実験における動物や人間の個体差，測定における測定者による差などがその例である。

そこで，実験の場においてもブロックと考えられ，その影響は制御因子や標示因子の効果の推定に偏りをもたらすという点でのみ関心の対象となる因子を，本来のブロック因子と呼び，それに対して応用の場においては再現性を持たないが実験の場においては再現性があり，かつそのばらつきの大きさ自体が推定すべきものとなっているようなものは（かりに）変動因子と呼ぶことにしよう。そうすると本来のブロック因子の効果については母数模型も変量模型もともに考慮することができるが，変動因子については本来変量模型を想定しなければならない。

また本来のブロック因子については，交互作用を想定することは無意味である．もしブロック因子と制御因子や標示因子との間に一定の交互作用があるとすれば，実験結果を応用の場に適用することができないことになって，因子効果の推定が無意味になってしまう。またブロック効果について変量模型を考えれば，交互作用を考えることも可能なように思われるかもしれないが，その場合にはブロック効果を主効果と交互作用に分解して考えることが無意味であって，それらはまとめて，いわば「一次誤差」として理解するのが適当である．現実にはブロック因子と他因子との間に交互作用に当るものが存在する可能性もあるが，そのような場合には割りつけのランダム化によって，それを誤差の中にふくめてしまうようにすることが必要である．

また2つ以上のブロック因子が存在する場合，すなわち2方以上の制約がある配置において，異なる因子相互間に交互作用があれば，それは本来方向性を考えること自体が無意味であったことを意味する

しかし上記のことはブロック因子と他因子との交互作用を検定することが，つねに無意味であるということを意味するものではない。ブロック設定の妥当性のチェックのために，そのような検定が必要になる場合もある．

2因子実験において，制御因子×制御因子の場合と，制御因子×標示因子の場合との最も大きな差は，前者は最適な水準の組み合わせを求めることが目標となるのに対して，後者では標示因子の各水準に対応して，制御因子の最適水準を求めることが問題となる点にある．そのことから，交互作用の意味の違いが生ずる．すなわち前者においては，交互作用が存在しないということは，制御因子のそれぞれの最適水準の組み合わせが，全体としての最適を与えるのに対して，後者では，制御因子の最適水準が，標示因子の水準によって変らないということを意味する。従って前者においては，実際には2つの因子の最適な水準の組み合わせの周囲についての交互作用のみが重要な意味を持つのに対して，後者については標示因子の各水準に対して，最適な制御因子の水準の近くで交互作用が存在するか否か重要である。

このようなことを考慮すれば，具体的な応用問題において単に交互作用の全体としての存在を検定するだけでは，実験結果の利用の目的に不十分であるといわねばならない，とくに両因子の水準数が多い場合には，交互作用の自由度が大きくなるから，総括的な，例えばF検定の検出力は，一般にはあまり大きくない，逆にまたF検定の結果が有意になっても，

交互作用が最適水準と無関係な方向にのみ存在しているのであれば、それは実はあまり意味がないかもしれない。交互作用の全体としての有意性だけでなく、その方向を知ることことが大切である。

従って直観的な方法としては、まず交互作用項の推定値を、水準の組み合わせに対応する観測値の平均から、2つの因子の主効果の推定値に対応する項を引いた"残差"と見て、それをプロットしてみることが必要である。くり返えしのある配置の場合には、更にこれと独立にその誤差標準偏差の推定量が得られるから、それを基準にして各項の有意性を見る必要がある。

いま X_{ijk} $i=1,\cdots p$ $j=1\cdots q$, $k=1,\cdots n$ を、因子の(i, j)水準に対応するk番目の観測値としよう。これについてここでは最も簡単な加法モデル：

$$X_{ijk}=\mu+\alpha_i+\beta_j+\gamma_{ij}+\varepsilon_{ijk} \qquad \varepsilon_{ijk}\sim N(0,\sigma^2)$$

を仮定しておく。そうすると

$$\hat{\alpha}_i=\bar{\bar{X}}_{i\cdot\cdot}-\bar{\bar{\bar{X}}}_{\cdots} \qquad \hat{\beta}_j=\bar{\bar{X}}_{\cdot j\cdot}-\bar{\bar{\bar{X}}}_{\cdots} \qquad \hat{\gamma}_{ij}=\bar{X}_{ij\cdot}-\bar{\bar{X}}_{i\cdot\cdot}-\bar{\bar{X}}_{\cdot j\cdot}+\bar{\bar{\bar{X}}}_{\cdots}$$

$$\hat{\sigma}^2=\sum\sum\sum(X_{ijk}-\bar{X}_{ij\cdot})^2/pq(n-1)$$

となることはいうまでもない。そうすると交互作用について単に平方和 $\sum\sum\hat{\gamma}_{ij}^2$ だけでなく、その個々の項の大きさを見る必要があり、その中で有意な大きさを示すのはどれとどれ

であるかを知る必要がある。そのためには各$\hat{\gamma}_{ij}$に対して、γ_{ij}の信頼区間の幅を示す必要がある。そのための簡単な方法は、各項毎に独立に求められる信頼区間、すなわちt分布を用いる、幅 $t_\alpha \hat{\sigma}/\sqrt{n(1-1/p)(1-1/q)}$ の区間と、すべての母数についての同時信頼区間、更にその最も簡単な方法としてScheffé の方式による幅 $\hat{\sigma}\sqrt{F_\alpha}/\sqrt{n(1-1/p)(1-1/q)}$ の区間、或いは$\sqrt{F_\alpha}$をより精密な値でおきかえたものとを、それぞれ与えることである。

具体的には、もし２つの因子がともに制御因子であれば、横軸に$\hat{\alpha}_i+\hat{\beta}_j$を、縦軸に$\hat{\gamma}_{ij}$を上記の幅をつけてプロットすることが必要である。一方が標示因子であれば各jごとに$\hat{\alpha}_i$の値に対応して$\hat{\gamma}_{ij}$をプロットすればよい。

このような問題をより形式的に定義すれば、結局γ_{ij}に関する多重決定問題と考えられるが、しかしここで$\sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$という条件があるから、それぞれを別個の母数として考えることはできない。さかのぼって考えれば

$$\mu_{ij} = E(x_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

とおくとき、母数

$$\delta(i_1 i_2, j_1 j_2) = \mu_{i_1 j_1} + \mu_{i_2 j_2} - \mu_{i_1 j_2} - \mu_{i_2 j_1}$$

を問題にすることになる。そうしてこれらの${}_pC_2 \times {}_qC_2$個の互いに独立でない母数の中で、どれとどれが有意に0と異な

るかを知ることが必要である。そのためには $\delta(i_1\ i_2,\ j_1\ j_2)$ に対して、Scheffé の方法による同時信頼区間すなわち

$$\bar{X}_{i_1j_1}+\bar{X}_{i_2j_2}-\bar{X}_{i_1j_2}-\bar{X}_{i_2j_1}\pm\hat{\sigma}\sqrt{4F_\alpha}/\sqrt{n}$$

を求めて、それが0をふくまない場合に母数が0と異なるものとすればよい。このような解析の結果は、甚だこみ入ったものになるから、それを簡単に表示することは難しいが、2因子がともに制御因子ならば、すべての $(i_1\ j_1)$ と $(i_2,\ j_2)$ の組み合わせに対して上記の δ をプロットすることが必要である。これに対して一方が標示因子の場合には、むしろ各 j ごとに

$$\delta^*(i_1,\ i_2,\ j)=\frac{1}{g-1}\sum_{j_2\neq j}\delta(i_1\ i_2\ j,\ j_2)$$

を定義し、

$\delta^*(i_1\ i_2\ j)$ に対する同時信頼域を求めて、各 j ごとにすべての $(i_1\ i_2)$ の組み合わせに対してそれをプロットすればよい。

またもし問題をより形式的に考えて、母平均の最大を与える因子水準を求めることと定義すれば、問題は、2制御因子の場合には

$$\mu_{i^*j^*}=\max_{i,j}\ \mu_{ij}$$

となる (i^*j^*) を求めること、1制御因子+1標示因子の場

合には，各 j に対して

$$\mu_{i^*(j),j} = \max_i \mu_{ij}$$

となる $i^*(j)$ を求めることである．

この場合，交互作用の有無を考えることは，結局前者については，

$$\bar{X}_{i^\circ j^\circ} = \max_{i,j} \bar{X}_{ij}.$$

となる (i°, j°) と，

$$\bar{\bar{X}}_{i^\Delta\cdot} = \max_i \bar{\bar{X}}_{i\cdot\cdot} \qquad \bar{\bar{X}}_{\cdot j^\Delta\cdot} = \max_j \bar{\bar{X}}_{\cdot j\cdot}$$

となる $(i^\Delta\ j^\Delta)$ のいずれをえらぶかという問題に，後者については，各 j ごとに一様に i^Δ をえらぶか，

$$\bar{X}_{i^\circ_j j\cdot} = \max_j \bar{X}_{ij\cdot}$$

となる $i^\circ(j)$ をえらぶかという問題になる．そうしてこのようにしてえらばれた水準を $(\hat{i}^*, \hat{j}^*)$ 或いは $\hat{i}^*(j)$ と表すとき損失関数として，前者については $\mu_{i^*j^*} - \mu_{\hat{i}^*\hat{j}^*}$ を，後者については $\sum_j (\mu_{i^*(j),j} - \mu_{\hat{i}^*(j),j})$ をとれば決定問題として定式化することができる．この場合交互作用項についての F 検定にもとづいて，2つの可能な候補をえらぶという方式は，明らかにあまりよい決定ルールではなさそうである．

いま2制御因子の場合について考えれば，このことは

$$\hat{\mu}_{ij} = \bar{\bar{X}} + \hat{\alpha}_i + \hat{\beta}_j + \phi(F)\,\hat{\gamma}_{ij}$$

ただし $\phi(F) = 1 \qquad F = \sum \hat{\gamma}_{ij}^2 / \hat{\sigma}^2 (p-1)(q-1) > F_\alpha$

$$= 0 \qquad F < F_\alpha$$

とおき，$\hat{\mu}_{i^*j^*} = \max_{i,j} \hat{\mu}_{ij}$ となる $(i^* j^*)$ をえらぶことに帰着する．ここで多母数推定に関する Stein 推定量を思い起こせば，上記の $\hat{\mu}_{ij}$ の代りに，例えば $\nu = (p-1)(q-1)$ とし，

$$\hat{\mu}_{ij} = \bar{\bar{\bar{X}}}_{\cdots} + \hat{\alpha}_{i\cdot} + \hat{\beta}_{j\cdot} + \max\left(0 \quad 1 - \frac{\nu - 2}{F}\right)\hat{\gamma}_{ij}$$

とすることが考えられる．

1 標示因子をふくむ場合には $p \geqq 4$ のとき

$$\hat{\alpha}_{i(j)} = \hat{\alpha}_i + \max\left(0, \quad 1 - \frac{p-3}{F}\right)\hat{\gamma}_{ij}$$

として $\hat{\alpha}_{i^*(j)} = \max_i \hat{\alpha}_{i(j)}$ によって $\hat{i}^*(j)$ をえらべばよいであろう．

しかしこのような方法は，なお検定統計量としての F 値にもとづいたものであるという点で，疑問が残る．もっと別のルールも考えられるが，ここでは単にどのようなルールにおいても制御因子と標示因子の区別が重要であることを強調しておきたい．

上記のような問題は，もっと複雑な構造を持つ実験に関して，いろいろと考えられるであろう．そうして多くの場合，単に特定の要因効果を検定して，それが有意であればその効

果を推定するというだけでは、データの解析として不十分であり、またその結論は必ずしもrelevantでないであろう。実験の具体的な目的を明確にすると同時に、各因子の性格ないし種類を正しく認識し、それを推測の方式に結びつけることが必要である。